BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : G06F 17/30 | A1

(11) International Publication Number: WO 00/62204

(43) International Publication Date: 19 October 2000 (19.10.00)

(21) International Application Number: PCT/US00/09767

(22) International Filing Date: 13 April 2000 (13.04.00)

(30) Priority Data:
60/129,104 14 April 1999 (14.04.99) US

(71) Applicant: CONJOIN, INC. [US/US]; Suite 355, 20 Mall Road, Burlington, MA 01803 (US).

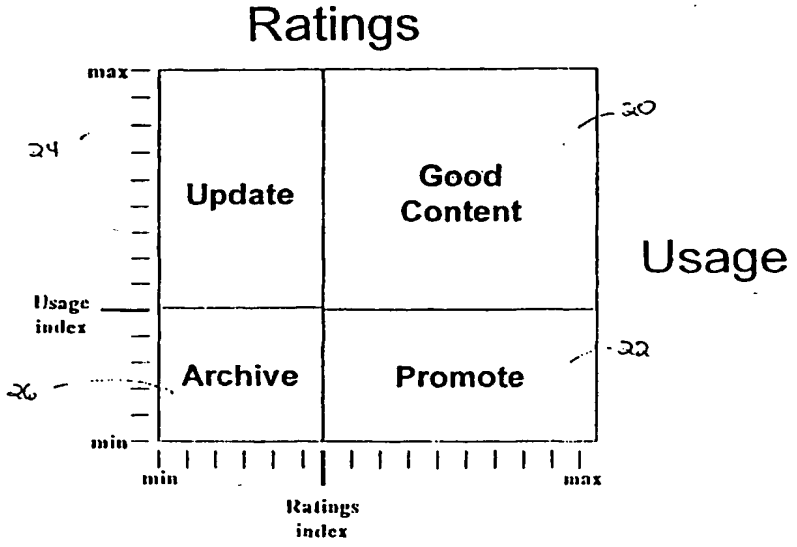
(72) Inventors: D'ARBELOFF, Nicholas; 345 Cross Street, Belmont, MA 02178 (US). DIMARE, Joseph; 197 High Street #12, Andover, MA 01810 (US). HEATH, Barbara; 28 Village View Road, Westford, MA 10886 (US).

(74) Agents: NUGENT, Elizabeth, E. et al.; Choate, Hall & Stewart, Exchange Place, 53 State Street, Boston, MA 02109 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ONLINE CONTENT LIFECYCLE MANAGEMENT BASED ON USAGE STATISTICS, USER-SUPPLIED VALUE RATINGS AND EXPIRATION DATES



(57) Abstract

Content Lifecycle Management (CLM) is a method for keeping content within an intranet (or internet) up–to–date and relevant for users of the website by archiving low–value, unused, inactive, and obsolete content from the site based on a combination of ratings threshold, usage threshold, and expiration date. Archiving content removes general access to that content, but retains it in the site for searches specific to the archive. CLM uses these threshold values as a method for providing automated feedback and notification to the publisher and/or author of aging or poorly performing content.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

ONLINE CONTENT LIFECYCLE MANAGEMENT BASED ON USAGE STATISTICS, USER-SUPPLIED VALUE RATINGS AND EXPIRATION DATES

Field of the Invention

The invention pertains to a system for maintaining and accessing a diverse collection of information to identify and keep information that is current and useful.

Background of the Invention

As information databases become more prevalent and more complete, a number of content management issues must be confronted. In particular, valuable data often become "lost in the haystack" of lower-quality content. In addition, systems rarely remove stale content or provide meaningful feedback to authors or publishers of the value of their content. It is an object of the present invention to address these issues by providing methods for winnowing databases to preserve high-quality content while archiving or deleting low-quality or stale content.

Summary of the Invention

In one aspect, the invention pertains to a system for maintaining currency and value for a collection of information. When a piece of information (a content item) is added to the collection, metadata is added to a database specifying the location of the information, and threshold values for usage and utility of the information. When the database item is accessed to access the information, the user is allowed to specify a value rating for the content item, and an access log is updated. The frequency of access of the content item and its value can thus be monitored to determine if either has fallen below the associated threshold value. If either usage or value rating falls below its threshold, the system performs some action, such as warning the author of the information that it is not being used or is received low ratings, marking the item for archiving, archiving the item, or deleting the item from the system. Content items may further have associated expiration dates which are given at the time that the content item is added to the collection. If a content item is approaching its expiration

date, the system may notify the author or perform another action such as archiving the item. The system can be used for monitoring diverse collections of information, including such items as URLs and web pages, files stored on a fileserver, whole directories and subdirectories, and nondigital data such as books.

Brief Description of the Drawing

The invention is described with reference to the several figures of the drawing, in which,

Figure 1 shows a sample screen containing content and a ratings entry area; and

Figure 2 shows the content matrix.

Detailed Description

Websites today are faced with daunting content management issues including the proliferation of old and stale content and poor capabilities of providing feedback and notification to the author. Content Lifecycle Management (CLM) is designed to eliminate this proliferation by keeping content that is displayed on an intranet site relevant and up-to-date. This goal is accomplished by tagging each item of content recorded in the content management database with an expiration date and threshold values. Threshold values are minimally acceptable levels in order for content to remain current and available. CLM uses at least 2 thresholds: one for the rating (value of the content) and one for the usage (how often each item is viewed.)

When content is published to the intranet, it is tagged with content thresholds for ratings and usage which the author or publisher feels represent minimum values. When content falls below either value based on a given timed period (for example, weekly basis), the CLM will notify the author of this fact, for example by e-mail, from within the CLM application, or both. The tabulated values which are required for this are collected from users who view and rate the content on the site. **Figure 1** is a screen shot illustrating how feedback is collected. Each time a user views a document (or item of content) it is registered in the usage table. When a user rates the content using the ratings engine **10** at the top of the window, it is also recorded

associating the content ID with the user ID, date, and rating. Rating of content may be either optional or required for the user.

The maximum and minimum values of the two thresholds form the CLM matrix, shown in **Figure 2**. Based on the rating and usage values, each content item can receive one of four classifications: (1) active or good content **20**, which means that this item of content is receiving high usage and that users find it of high value; (2) needs promotion **22**, which means that users value the content but usage is low (usually because users can not easily find the content); (3) needs updating **24**, which means that users access the file frequently but that it is not of significant use or value; and (4) archive content **26**, which is the default action for content that is not being viewed often and is of low value.

The ratings and usage of content in combination with its expiration date serve as an effective indicator for archiving content.

The inventive system for maintaining currency of the information dovetails well with systems for targeting information to particular users or groups of users. For example, the methods described herein may be used with the methods described in U.S. Provisional Application No. 60/129,106, filed April 13, 1999, and U.S. Patent Application "Group Targeted Content Personalization," (attorney docket 2001774-0001), filed on even date herewith.

Example

In one embodiment of the invention, Content Lifecycle Management is based on a set of database tables that are part of an overall intranet database. The following data are cataloged in the database tables:

- Content information – may include ID, name, file name, content type, publish and expiration dates, author/publisher, approval and archive status, and ratings/usage thresholds for each content item
- Ratings – may include content ID, user ID, date, rating, and module in which the content was rated
- Usage - may include content ID, user ID, date, and module in which the content was viewed

• Author/Publisher Info – may include e-mail address for notification by e-mail

The following data are stored on durable media such as a hard disk:

• Content, which may include documents, files, data, executables, and catalogue entries or other references to nondigital content.

• Archiving Daemon (code)

• Low-Usage Daemon (code)

• Low-Rating Daemon (code)

• CLM Matrix Daemon (code)

Content Publishing

Content is published to the intranet using the publishing functionality within the system. During the publishing process, the file may be uploaded and stored on the server and all file-specific and publisher-entered information as well as the association between the file and the database entry is saved in the database tables. Data specific content such as web links (URLs) are stored in the database and the information is associated with the database entry. The user-entered information may include items such as the classification of the document using the topic and subtopic parameters, access classification (who has access to view the document), target classification (who would most benefit from this content), expiration date and the rating and usage thresholds for the content. (Access and target classification are discussed more fully in U.S. Provisional Application No. 60/129,106, filed April 13, 1999, and U.S. Patent Application "Group Targeted Content Personalization," (attorney docket 2001774-0001), filed on even date herewith).

Archiving Daemon

The Archiving Daemon comprises executable code which is run as a scheduled task on a periodic basis (set by the administrator). In one preferred embodiment, it is set to run on a nightly basis. Each time the archiving daemon is run, it makes two comparisons. First, it compares the expiration date of each record in the content library with the current date. If the current date matches the expiration

date, the record is archived and an e-mail is sent to the author notifying the author that the record has been archived. At this point, the content that has been archived is no longer available throughout the site unless a search on archived content is executed.

If the current date represents a period of a week before the expiration date (or any other suitable time period), the archiving daemon sends an alert e-mail to notify the author that the content is aging and will be retired at the end of the week unless it is updated.

Low-Usage & Low-Rating Daemons

These daemons are used to notify the author if content the author has published falls below the rating and usage thresholds set when the content was published. Like the Archiving Daemon, they may be set to run on a periodic basis (*e.g.*, nightly). These Daemons may further directly archive content which falls below one of the thresholds (*e.g.*, if the author/publisher does not respond to an invitation to improve the content).

CFM Matrix Daemon

The Content Lifecycle Management Daemon compares the averaged ratings and the usage values of each record in the content management database against the administration-set values for both the rating and usage. The administration set values are the *index values* and allow the matrix to be configured to a particular group of users. If the rating and usage values fall above the index values for both ratings and usage, no notification is sent. If either the usage or the ratings values fall below their respective index value, a notification is sent.

Other embodiments of the invention will be apparent to those skilled in the art from a consideration of the specification or practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with the true scope and spirit of the invention being indicated by the following claims.

What is claimed is:

1. A computer-implemented method of maintaining currency and value of a collection of information, comprising:

adding an item corresponding to a single piece of information to a database of metadata pertaining to the collection of information, where the added item includes

an identifier indicating a location for the piece of information;

a minimum value level for the piece of information; and

a minimum usage level for the piece of information;

accessing the added item in the database of metadata, where accessing includes

using the identifier to access the piece of information;

updating a record of the actual usage level for the piece of information; and

allowing updating of a record of the actual value rating for the piece of information; and

performing an action in response to a condition in which

the actual value rating is below the minimum value level; or

the actual usage level is below the minimum usage level.

2. The method of claim 1, wherein the item further includes an expiration date for the piece of information, and wherein an action is performed in response to a condition in which the actual date is within a selected time period from or is equal to the expiration date.

3. The method of claim 1 or 2, wherein the action is selected from the group consisting of:

removing the item from the database of metadata;

marking the item for archiving;

placing the item in an archive database; and

notifying a user that the condition exists.

4. The method of claim 1, wherein the identifier is selected from the group consisting of uniform resource locators, file location paths, directories, subdirectories, and catalogue entries for nondigital information.

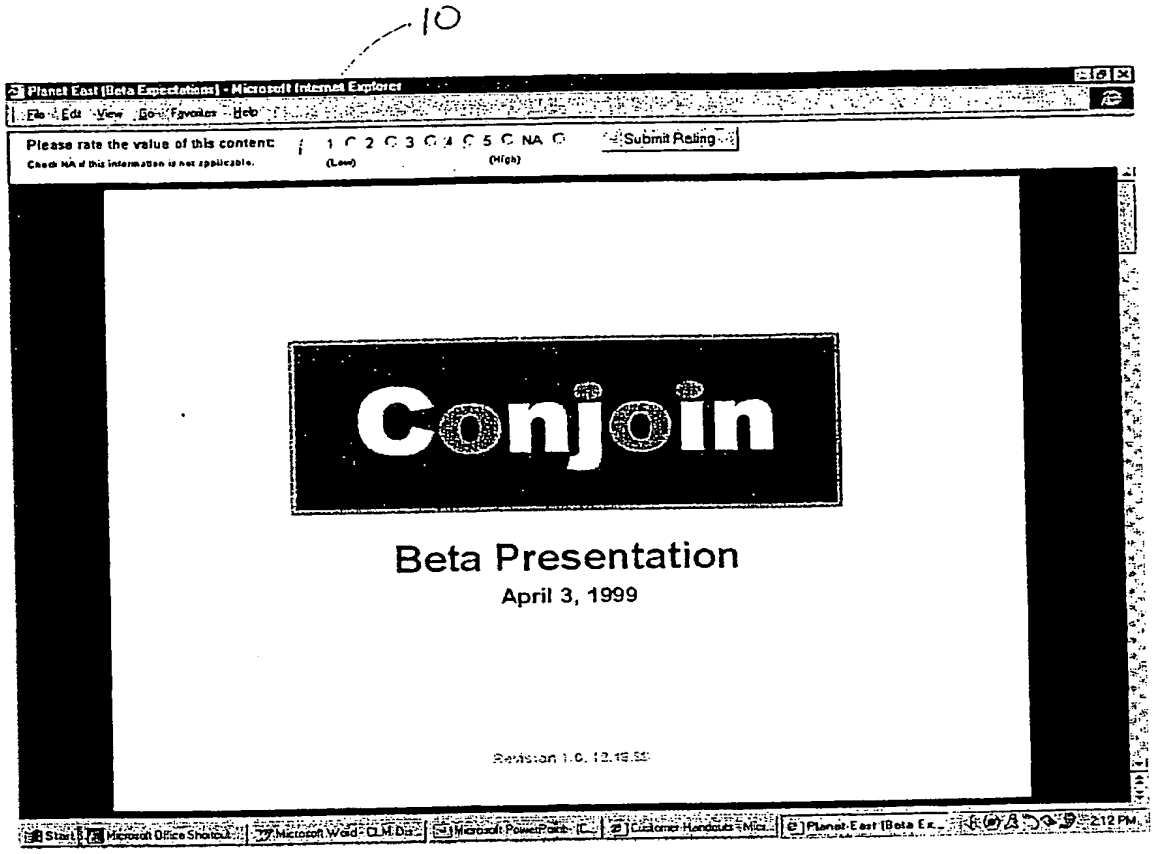

10
Planet East [Beta Expectations] - Microsoft Internet Explorer
File Edit View Go Favorites Help
Please rate the value of this content:
Check NA if this information is not applicable.
1 2 3 4 5 NA
(Low)
(High)
Submit Rating
Conjoin
Beta Presentation
April 3, 1999
Revision 1.0, 12.15.98
Start
Microsoft Word - CLM Dis...
Microsoft PowerPoint - [...
Customer Handouts - Mic...
Planet East [Beta Ex...
2:12 PM

Figure 1

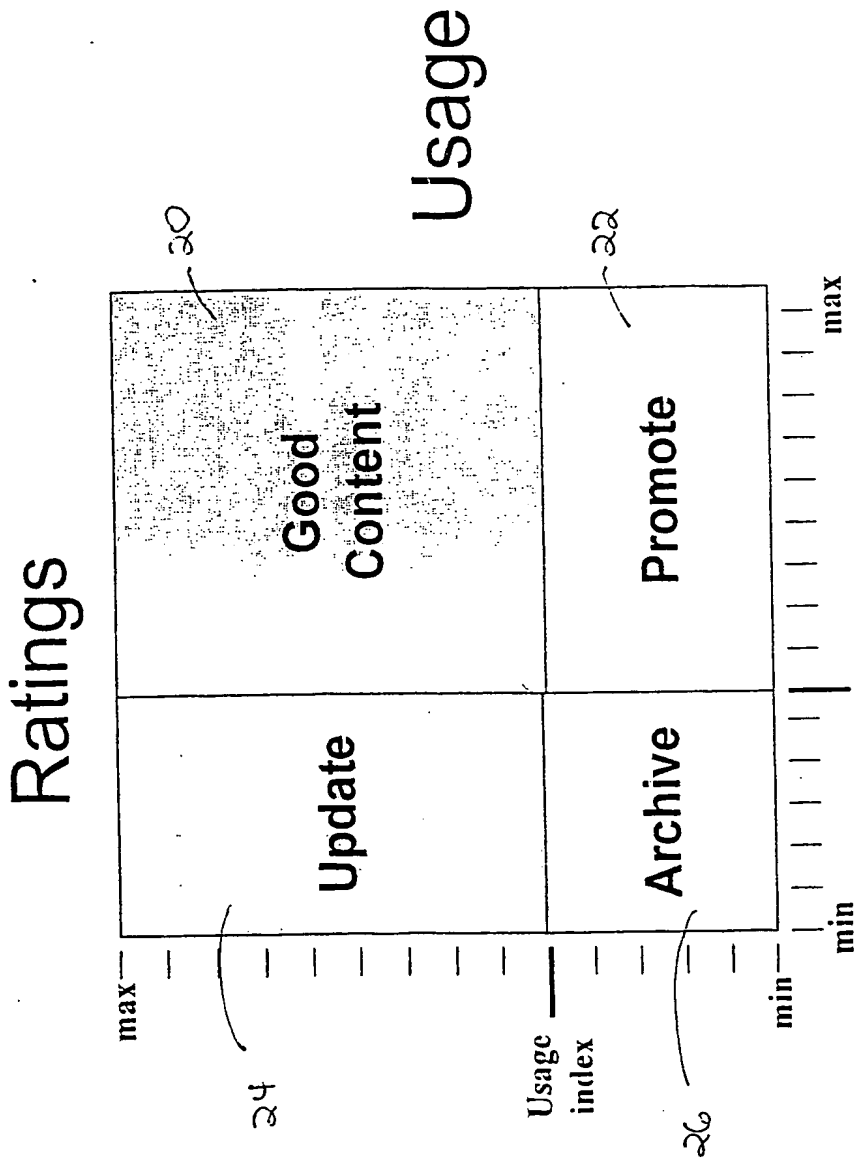

Usage
Ratings
Good Content
Promote
Update
Archive
max
min
max
min
Usage index
Ratings index
30
32
24
26

Figure 2

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 00/09767

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
WPI Data, EPO-Internal, PAJ, INSPEC, COMPENDEX, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 515 073 A (HEWLETT PACKARD CO) 25 November 1992 (1992-11-25) abstract column 2, line 23 - line 48 column 5, line 24 -column 9, line 32; figures 2-4	1,3,4
A	GB 2 327 787 A (KNOWLEDGE HORIZONS PTY LTD) 3 February 1999 (1999-02-03) abstract page 5, line 21 -page 6, line 18 page 7, line 30 -page 8, line 9	1,2,4

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 August 2000	30/08/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Polzer, A

3

Form PCT/ISA/210 (second sheet) (July 1992)

SDOCID: <WO___0062204A1_I_>

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int tional Application No
PCT/US 00/09767

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0515073 A	25-11-1992	US 5313631 A	17-05-1994
		DE 69224678 D	16-04-1998
		DE 69224678 T	02-07-1998
		JP 5158770 A	25-06-1993
GB 2327787 A	03-02-1999	AU 6905498 A	03-12-1998